**ZIF Documentation**

by [Fowler, M. (Mark)](http://vsnsbiovdc.ent.dfo-mpo.ca/author/fowler) — last modified Apr 19, 2012 01:07 PM — [History](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures/@@historyview)

**Description of the processing, coding and formatting of ZIF data since 1986.**

**CONTENTS**

1. [Introduction](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-0)
2. [SECTION I â General](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-1)
3. [SECTION II â Historical ZIF: 1986-2002](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-2)
   1. [Maritimes:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-3)
   2. [Newfoundland:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-4)
   3. [Quebec:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-5)
   4. [Gulf:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-6)
   5. [All Regions:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-7)
4. [SECTION III - Interim Years: 2003-2006](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-8)
   1. [Maritimes:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-9)
   2. [Newfoundland:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-10)
   3. [Quebec:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-11)
   4. [Gulf:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-12)
   5. [All Regions:](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-13)
5. [SECTION IV â ZIF2: 2007->](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-14)
6. [APPENDICES - Items linked to Sections I-IV](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-15)
7. [Details on Assorted Subjects](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-16)
   1. [About 1994](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-17)
   2. [Subtrips](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-18)
   3. [Example](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-19)
   4. [log and slip tables](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-20)
   5. [processing](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-21)
   6. [Recent Documentation Review Notes](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-22)
8. [Creation of a Continuous ZIF Database](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-23)
9. [Project List (noted if done)](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-24)
   1. [other regions](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-25)
   2. [historical ZIF effort variables](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-26)
   3. [species codes (done)](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-27)
   4. [unspecified](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-28)
   5. [unknown depths](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-29)
10. [Minor Task List (noted if done)](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-30)
    1. [effort\_flag (done)](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-31)
    2. [confusion (done)](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-32)
    3. [main species caught](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-33)
    4. [latitude/longitude (done)](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/status-1986-2002-zif-procedures#section-34)

**Introduction**

The Zonal Interchange Formatted File (ZIFF, typically shortened to ZIF) concept was developed as a means of merging and sharing commercial fisheries data between regions. This facilitated research and stock assessment on species/populations that crossed regional boundaries, and also accounted for the landing of catches outside the region of capture. To accomplish this a set of variables were identified for inclusion, and common field formats were designed to hold the contents of these variables.

It is sometimes assumed that because ZIF data is represented in single annual files of common format that this implies equivalence in the substance of the data between regions. This may have been an objective, but if so it was never achieved. Do not equate commonality in the multiregional ZIF field format with commonality in underlying data structure, coding, or content. Set aggregation protocols vary between regions, and within region between years. A given variable may not be populated (blank or filled with a single null-proxy value) by a given region. And various coding systems evolved independently in each region, often from a common FAO/STACAC origin which can give a false impression of common coding. Coding differences between regions for [**Gear codes**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) and [**Species codes**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) have been investigated. These codes are standardized in the current database, which could produce different results compared with older databases.

Much of the historical documentation on ZIF that has been available to Population Ecology Division pertains to the Maritimes Region data, not that of the other Atlantic regions (Newfoundland, Gulf of St Lawrence, Quebec). We have a reasonable handle on changes over time with respect to Maritimes Region data, much less so for the [other regions](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#other-regions).

The ZIF documentation presented here is partitioned into four main sections delimited by time periods that reflect the most pervasive changes in the structure, contents and formats of ZIF data. The first section contains subject matter that encompasses all time periods (it is a very small section). The second section focuses on the 1986-2002 period, encapsulating the data familiar to past users of ZIF. The third section covers the period 2003-2006, labelled the Interim Years, during which ZIF data was generally unavailable to DFO Science. An ad hoc form of ZIF, but not imposing ZIF formats and contents, was maintained at National Headquarters, and we have processed this into a structure/format consistent with later years. The fourth section covers the ZIF since 2007, when it was redesigned.

The revised ZIF starting in 2007, often referred to (unofficially) as ZIF2, was a rationalization exercise motivated by Policy & Economics Branch, primarily to improve the quality of the economic data associated with commercial fisheries. Fortunately DFO Science was included in the project (unfortunately not in every region), so a number of improvements were affected of relevance to fisheries science, aside from simply being able to obtain any ZIF at all.

**SECTION I â General**

There is little of substance about ZIF that does not require qualification with respect to time. Even some of the contents of this section have temporal aspects, they just arenât coherently documented (subjects essentially specific to other regions). Had to throw them somewhere.

Null catches are eliminated in ZIF. And catches are rounded to the nearest kilogram, resulting in false zeroes for catches under .5 kg.

Within a given year the most recent tonnage and length data for a CFV are applied to the whole year, of relevance if either changed during the year.

The nature of the vessel information associated with a given CFV varies between regions with respect to vessels landing in a region other than their own.

The TWO\_HUND\_MILE variable was only filled by Newfoundland (and only in some years).

The GULF\_BASED\_VESSEL flag was usually only filled by Gulf Region, and exhibits inconsistent coding over time. It served different purposes over time, but we donât know the details [see [other regions](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#other-regions)].

**SECTION II â Historical ZIF: 1986-2002**

**Maritimes:**

ZIF data was fully aggregated to what are called [subtrips](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/" \l "subtrips" \o "Subtrips) prior to 1986, while 1986-1998 is a mixed bag of changing and largely undocumented processing protocols, [with 1994](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#about-1994)  an extreme example of processing errors due to bureaucratic problems. Since 1995 the Maritmes region ZIF records can be set-specific if the log distinguished sets, but the rules to determine where this was done are unknown. Also the effort was still aggregate-processed with the subtrip results repeated across the set records.  Subtrip CPUE was obliged to use ZIF to analyze effort time series, remembering to break on the effort data values but sum the catch data, and be aware that 1986-1995 will be variously inconsistent with later years and within themselves. A suite of [historical ZIF effort variables](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#historical-zif-effort-variables) comprising Days\_Sea, Days\_Ground, and Days\_Fished, provides a somewhat dubious but global estimate of effort regardless of source data type, but you may have difficulty pinning down how it was calculated for given gears and years, plus it derives from trip, not subtrip, aggregations (and again then repeated across all records for the trip, so break, don't sum).

Retention of set data during the 1996-1998 period varies. If cod and haddock were part of the catch, set records were often kept unaggregated. In the absence of cod or haddock, previous aggregate processing protocols tended to be carried through. Data associated with cod or haddock also was often audited and edited, while the rest of the groundfish data was not.

Theoretically, ZIF data since 1999 was not to be aggregated, the intent being to preserve the set-specific data where provided. Unfortunately this was not truly accomplished. The original log records were represented in ZIF with respect to the actual catch, such that you could match species by weight back to log records. However subtrip processing was still applied to the effort data and geographic coordinates (latitude and longitude). This had the result of repeating subtrip effort rollups across all the individual sets, worse than using subtrip-aggregated data in the first place [see [example](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#example)]. Each record repeats the subtrip-summed effort data, and the subtrip-determined lat and long (usually that of the first set of a subtrip). And as we had no unique set identifier, we had no easy way within the ZIF dataset itself to determine whether a given record was a rollup or a set, since both types of data occur (e.g. slip-only source data).

Catch dates do not start showing up until 1988. Latitude/longitude coordinates do not start showing up until 1990.

Catch dates for trips spanning two years were âeditedâ to match the titular year of the ZIF file. Thus January catch dates from trips begun in December were coerced to be December 31 of the preceding year. Regard a date of December 31 with suspicion.

The definition of [MAIN SPECIES CAUGHT](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#main-species-caught) has been that of the biggest catch by weight, with the singular exception of halibut, which got the main species tag if it dominated in terms of value. This exception was cancelled in 2002, with the inception of MARFIS. Since that time we have just the one definition across the boards. Possibly other deviations existed?

Species codes 160 and 161 only occurred in 1999 of the 1986-2002 ZIF years, and equal species codes 356 (skate) and 362 (dogfish). This has caused [confusion](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#-confusion) in the past, especially as Newfoundland switched from 356/362 to 160/161 in 1998, the Gulf switched in 1999, and Maritimes and Quebec regions switched from 2003.

Many of the processing issues that concern Maritimes Region data discussed above (end-of-year catch dates and set aggregation protocols) and in the General subsection below (longline hooks, depth data, species identification) can be resolved using the [log and slip tables](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#log-and-slip-tables) from which the Maritimes ZIF data component was derived. The task is not seamless, as the ZIF-processed data was subject to editing and processing that was not applied to the source log and slip data. Thus a systematic program will have a bunch of non-matching records between data sources. And 1994, as mentioned earlier, is very problematic, such that 1995 became a common starting year for CPUE series.

**Newfoundland:**

Never aggregated data during processing, so if the log data distinguished sets these are individual records in ZIF, all the way back to 1986. Set series can often be discerned by large numbers of records for a given species and boat on one day, but the data offers no systematic means of telling an aggregate record from a set. Even summary log records can result in multiple occurrences (3-4) on one day, with no definitive way to tell if set or summary data (one might reason it out to varying extents based on numbers of records and effort values). Numbers of hooks are the original counts, no coding system used. Latitude/longitude coordinates start showing up in 1994.

**Quebec:**

May have used daily summary logs throughout, as they do today. Numbers of hooks associated with longline gears have always been divided by 1000 and rounded to produce a coded effort scale. Catch dates start showing up in 1987, not becoming common until 1988. Latitude/longitude coordinates start showing up in 1991. The minutes were recorded in 10âs through 1998, giving the data a less continuous pattern than other regions. Since 1999 both digits are accounted for.

**Gulf:**

May have had the same or similar aggregation protocols as Maritimes Region until 1999. From 1999 onward Gulf Region continued to aggregeate sets. Numbers of hooks associated with longline gears were not included in their data, with the singular exception of 2000. In 2000 the number of hooks were provided in coded form (divided by 1000 and rounded to produce a coded effort scale). Latitude/longitude coordinates start showing up in 1991.

**All Regions:**

The gear codes are contained in the gear\_type variable during this period, not the gear\_code variable as some might expect. The gear\_code variable was only used by Quebec during 1991-1998 to hold specifics related to dredge gears in their region. The variable is variously blank or 0-filled by all other regions, and by Quebec for all other gears (and years before 1991 or after 1998)

 Regional [species codes](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#-species-codes) diverged over time. This obliged a rationalization across regions to produce the standardized ZIF2, but this did not address earlier years.  [**Species Codes**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) lists the current codes and species names. It also includes comments related to regional differences in coding and introduces several new codes as alternatives to various âunspecifiedâ constraints on past codes.

 As each region had different protocols for processing the number hooks associated with longline gears, and these changed over time within each region as well, a table ([**Data Validation - Longline Hooks**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation)) summarizes the protocols by region and year.

For most regions, there is no way to distinguish whether a depth zone of 0 was originally an [unknown depth](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#unknown-depths) or the code for a depth over 250f. The depth zone field does not have blanks and 0's as indicated by the old dictionary. The only exception is Gulf Region, where blanks and 0âs are used as intended in every year except 2000, where blanks are zero-filled like all other regions. Also note that Maritmes region did not start processing depth data until 1988 (all records 0 in 1986-1987). Newfoundland began using the defined blank/zero protocol starting in 1995, but it was only applied partially, such that both blanks and 0âs denote unknown depths, while 0 also denotes depths over 250f. It appears Newfoundland began trying to fix the problem in 1998, with the unfortunate result of 1-filling the variable in 1998-1999, after which they put unknown depths under code 1 (perhaps reasoning that preserving the integrity of the very deep code was more important than that of the shallowest sets (see [other regions](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#other-regions)).  The regional/yearly development of depth data processing is summarized in [**Data Validation - Murky Depths**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation). The depth codes are standardized during post-load processing to produce the current ZIF database, but this cannot recover the depth information that was lost by the original miscoding.

 Some species were default coded as â[Unspecified](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#unspecified) \_\_\_â by DMP-related keypunchers, regardless of whether species were identified or not. This was very common with flatfish species. Regional keys were used to prorate flatfish catches for 1961-1990, any species-specific log or DMP information being generally disregarded.This protocol was found to be inappropriate, and since 1991 attempts have been made to achieve identification. However the change in protocol has not been consistently applied in practice, even when the log data is species-specific. The ZIF data since 1991 variably identifies 40-70% of the flatfish landings to species. Reconciliation of ZIF and log data achieves 60-90% species-specific catch data.

Species form. These codes determine the live weight conversion factors applied to a processed catch. They differ between regions, and within regions over time as well. And codes with the same form definition in two or more regions may be associated with different live weight conversion factors (see SECTION IV). It should be noted that a change in conversion factors need not imply that a previous conversion factor is incorrect. They can change over time due to changes in growth (e.g. fat years, lean years). Thus it would not be appropriate to apply a later conversion factor to previous years by default - we would have to confirm the applicability by year. We have little knowledge of these codes for other regions during the 1986-2002 period.

Historical documentation suggests that dates from trips spanning two years were changed to make records âfitâ to the titular year of the annual ZIF files. All dates in the previous year were changed to the first day of the current year yyyy0101).  All dates in the following year were changed to the last day of the current year (yyyy1231).  As well, up until about 2002, if the CATCH DATE was missing it might be replaced with the LANDED DATE and vice versa. An examination of December-January date patterns by region suggests that Maritimes Region may have been the only practitioner of year-end adjustments to valid dates, and this may only have been applied to subsequent year catch dates, as evidenced by âspurtsâ of yyyy1231. The aggregation protocols used by Gulf and Quebec Regions might have masked similar processing (and make it a moot point anyway). And for Newfoundland region we sometimes encounter catch dates from the year preceding the titular year of ZIF, suggesting they did not revise dates. Possibly the notion of yyyy0101 conversions was a misunderstanding, as any candidates would equate to the yyyy1231 conversion of the preceding year.

**SECTION III - Interim Years: 2003-2006**

**Maritimes:**

ZIF data no longer confounded with subtrip processing (i.e. no recoding of numbers of hooks, no replacing set-specific effort data with subtrip values) and no aggregations are performed. The [effort\_flag](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/" \l "effort_flag" \o "effort_flag) variable tags a âYâ to one record (species) of each set to facilitate set-specific association with effort data to avoid redundancy, with all other records for the set marked with âNâ. And no editing of catch dates was conducted to force records to match the year of the ZIF file (a date of December 31 is no longer suspicious).

The definition of [MAIN SPECIES CAUGHT](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#main-species-caught) is that of the biggest catch by weight, no exception for halibut anymore.

In 2006 Maritmes used seqnum to hold LOG\_EFRT\_STD\_INFO\_ID for the first time, providing a sequential set (or subtrip if summary logging) number where applicable.

**Newfoundland:**

No aggregation during processing. If the log data distinguished sets these are individual records in ZIF. Sets can be identified by set numbers (seqnum) for every year except 2006, where seqnum is a sequential database record number (oversight, might be fixed). The [effort\_flag](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/" \l "effort_flag" \o "effort_flag) variable tags a Y (2003,2006) or â (2004,2005) to one record (species) of each set to facilitate set-specific association with effort data to avoid redundancy, all other records marked with âNâ or a blank. Numbers of hooks are the original counts, no coding system used.

**Quebec:**

Used daily summary logs, no set data being collected. The [effort\_flag](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/" \l "effort_flag" \o "effort_flag) variable tags a Y (2003,2006) or â (2004,2005) to one record (species) of each subtrip to facilitate association with effort data to avoid redundancy, all other records marked with âNâ or a blank. Numbers of hooks associated with longline gears divided by 1000 and rounded to produce a coded effort scale.

**Gulf:**

Sets were aggregated according to whatever subtrip definitions were applied, possibly equivalent to subtrip definitions in Maritimes region, with the possible exception of 2003 (might include set records). The [effort\_flag](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/" \l "effort_flag" \o "effort_flag) variable tags a Y (2003,2006) or â (2004,2005) to one record (species) of each subtrip to facilitate association with effort data to avoid redundancy, all other records marked with âNâ or a blank. Some of the records in 2003 have blank effort-flag values. Numbers of hooks associated with longline gears were provided in coded form (divided by 1000 and rounded to produce a coded effort scale).

Catch dates for trips spanning two years were âeditedâ to match the titular year of the ZIF file. Thus January catch dates from trips begun in December were coerced to be December 31 of the preceding year. Regard a date of December 31 with suspicion. This might cause some confusion, as this is how Maritimes Region processed catch dates from 1988 to 2002, but has left them alone ever since, while Gulf Region did not appear to change such catch dates in earlier years. However they might have had the same protocol earlier, it was just irrelevant given the aggregate processing of the data. Thus the sudden appearance of December 31 spurts might reflect more general changes in processing, unknown to us at present (see[other regions](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#other-regions)).

**All Regions:**

The compiled ZIF changed format every year since 2003, reflecting both the ad hoc nature of the process since 2003, plus evolution toward the âZIF2â format achieved for 2007. New variables appear and old variables disappear, although the distinction may be clouded by the VDC depiction of ZIF. We used to get a ZIF file from each region and process them ourselves to produce the various components of the database we associate with ZIF (CATCHERS\_yyyy, TRIPS\_yyyy, SUB\_TRIPS\_yyyy, IDENTIFIED\_CATCHES\_yyyy). The actual ZIF data was just a single ASCII file per region, formatted similarly to the currently defined âcore fieldsâ in the [**ZIF 2007 Definition**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation). Not all these variables may have appeared in the VDC ZIF database, plus we created some of our own. Adding to the confusion, a given year between 2003 and 2006 could include any number of region-specific or HQ-processing variables. It is also possible that some of variables denoted as core fields were not part of the 1986-2002 ZIF, but were provided since 2003, and the ZIF2-related documentation trail might not account for pre-2003 data very well. We processed the 2003-2006 files into the new 2007 format.

The definition of [MAIN SPECIES CAUGHT](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#main-species-caught) is assumed to be that of the biggest catch by weight.

[Depth zone](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#unknown-depths) continued to be problematic [see [**Data** **Validation-Murky Depths**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation)]. The Newfoundland Region continued coding unknown depths into depth zone 1, but provided the actual depth data itself, making the depth code irrelevant. Maritimes Region stopped providing depth data until 2006 (so nothing for 2003-2005), at which point they just provided the actual depths (no depth codes). Quebec provided no depth data in 2003, correctly coded depth zones in 2004-2005, and incorrectly coded depth zones (unknown=0) in 2006. Gulf region stopped providing depth data of any kind. The depth codes are standardized during post-load processing to produce the current ZIF database.

Overlapping CFVs between Gulf and Newfoundland regions. The region variable must be used to differentiate such vessels.

Species form. These codes determine the live weight conversion factors applied to a processed catch. They differ between regions, and within regions over time as well. And codes with the same form definition in two or more regions may be associated with different live weight conversion factors. It should be noted that a change in conversion factors need not imply that a previous conversion factor is incorrect. They can change over time due to changes in growth (e.g. fat years, lean years). Thus it would not be appropriate to apply a later conversion factor to previous years by default - we would have to confirm the applicability by year. The 2003-2006 years recieved some attention during a project to rationalize conversion factors across regions (see SECTION IV), the documentation for which may be of interest.

[Latitude/longitude](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#latitude-longitude) formats differ between regions since 2003. Maritimes and Newfoundland use decimal degrees, Gulf and Quebec put the actual minutes in the fractional portion (i.e. -64.55 means 64 degrees 55 minutes). These are standardized during post-load processing to produce the current ZIF database.

Substitutions for blank values between CATCH DATE and LANDED DATE no longer conducted.

Derivation of the MONTH and YEAR variables uses the historical convention of coercing post-end-of-year dates to December 31, but without editing the actual CATCH DATE and LANDED DATE.

Gear codes are contained in the gearcode variable (in earlier years it was kept in the gear\_type variable, while the gear\_code variable contained detailed dredge codes for Quebec Region).

**SECTION IV â ZIF2: 2007->**

The regional datasets are [merged](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#processing) in Ottawa, where differences in formatting and coding are reconciled. Generally a lot more data validation than before, and at the multiregional level, such that many variables that used to be associated with a lot of blanks will now be filled in. And the exercise to develop a better design resulted in several new variables being created, many of which started showing up over the course of 2003-2006. From a DFO Science perspective the most relevant new variables might be FIN (Fisher Identification Number), LICENCE, FISHING\_AREA (fishery-specific area codes), depth (in fathoms, not codes), CFV\_FLAG & CFV\_TYPE, FIN\_FLAG & FIN\_TYPE, and SEQNUM.

The CFV\_FLAG/TYPE variables clarify the CFV as that of a registered vessel or not, among other details. Similarly the FIN\_FLAG/TYPE variables clarify if the FIN is for a fisherman, a company, boat, etc. Both these pairs of variables are coded in detail by Maritimes Region, less so by other regions.

SEQNUM, in conjunction with EFFORT\_FLAG, provides an easier way to distinguish sets without having to process break variables or make as many assumptions about subtrip rollups as in the past. SEQNUM represents the sequential set number where applicable (for Maritimes Region it is the MARFIS variable LOG\_EFRT\_STD\_INFO\_ID), As Gulf and Quebec Regions do not address catch information to the level of a set, SEQNUM is blank for their data. [**Waiting for confirmation on nulls vs values logic for Maritimes and Newfoundland regions**]

Depth zone continued to be problematic ([**Data**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) [**Validation-Murky Depths**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation)**),** however every region but the Gulf provided actual depth data. Maritimes Region coded unknown depths into depth zone 0. Newfoundland and Quebec provided correctly coded depth zones. Gulf region did not provide depth data of any kind. Post-load processing used the actual depth data to solve the coding issue with Maritimes region.

A number of [species codes](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#-species-codes) differ between MARFIS (regional data) and the new ZIFF. [**Species Codes.xls**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) lists the current codes and species names in ZIF2. It also includes comments related to regional differences in coding and introduces several new codes as alternatives to various âunspecifiedâ constraints on past codes.

Be aware of some [species code](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#-species-codes) confusion pertaining to wollffish. The codes in Maritimes Region through 2008 were 174 (unspecified wolffish), possibly 185 (striped), possibly 186 (spotted), and 187 (northern), and we rarely saw anything but 174. Beginning in 2007 these were supposed to become 191 (northern), 192 (striped), 193 (spotted), and 194 (unspecified wolffish). However not everyone was aware that catfish=wolffish, such that 174 continued to predominate through 2008 from Quebec and Maritimes Regions.

Species form. These codes and the live weight conversion factors associated with them were investigated exensively during 2006-2007 to variously standardize them across regions or confirm differences are justified. A [**live weight conversion chart**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) is produced each year to accompany the annual ZIF update. It should be noted that a change in conversion factors need not imply that a previous conversion factor is incorrect. They can change over time due to changes in growth (e.g. fat years, lean years) or adjustment for the dominant size classes in a fishery. Thus it would not be appropriate to apply a new conversion factor to previous years by default - we would have to confirm the applicability by year.

A list of [**gear codes**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) developed during the ZIF2 exercise provides a more complete set of definitions than have been associated with local databases. Many of these gears will be specific to other regions, although we often see allegedly âextra-regionalâ codes in other regions data. For common gears, the merged list across regions is sufficient. For a few uncommon gears different regions may use the same code for different gears. The codes are standardized in the ZIFF since 2007, but not in earlier years. A time series snapshot of record frequencies by gear tries to provide an overview of gear codes with some historical and regional contrast ([**Data**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) [**Validation - Gears**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation)**)**. Sometimes the invalid code of 0 was used in place of unknown (code 99).

The [**ZIFF 2007 Definition**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) prepared by HQ to describe the new ZIF includes some instances of historical depictions. This background information is largely based on recent years. As the ZIF was not updated locally, past users of ZIF data would be familiar with 1986-2002 data, but not the 2003-2006 data upon which much of the explanatory text is based. And a lot of changes occurred when the regional ZIF update process ceased in 2003. Those considered most relevant to Science staff are addressed in the four main sections of this web doc. A few more variables that diverge from the descriptions in ZIFF 2007.doc have been reproduced with some [review notes](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#recent-documentation-review-notes) that some might find helpful.

**APPENDICES - Items linked to Sections I-IV**

**Details on Assorted Subjects**

**About 1994**

First Management ran short of DMP-related forms. Somewhat coincident with this, fixed gear was exempted from DMP in the 1994 Plan.  Then the replacement forms were to new designs, never seen by Science, that variously didn't allow for effort data or could only do trip summaries. So log submission became voluntary, and of variable utility depending on supplies of old vs new logs.Took a while for Science to become aware of the changes, and then Science was upset. Everything got put back to 1993 rules in 1995. And the above is a simplification (not entirely correct in details, which were more complex, and will include unknowns).

**Subtrips**

Up to 1999, the ZIF data processing usually aggregated log records into subtrips, where a subtrip was defined as at least equivalence in boat, unit area and date. It could also include breaks according to gear type, main species caught, and depth, but the rules lacked consistency[subtrips], and quite often the predominant gear and target species was determined at trip or subtrip levels prior to aggregation. It may have stabilized from 1995 as breaking on the variables for region, CFV, tonnage class (not sure why), gear, NAFO area, dates landed and caught, species sought, and depth (rarely to be had).

**Example**

 Simple example of Maritimes Region ZIFF set-specific ârollupâ: If an otter trawler did 3 sets at 3 different lat/longs but all within a subtrip definition, and say of 1, 2 and 3 hours duration each, we got 3 records as desired, and the correct catch per record. However all 3 records would have the same lat/long (from the first set) and each record would have a tow duration of 6 hours (the sum of all 3 sets).

**log and slip tables**

The format of lat/long fields in log\_detail fields was changed to decimal degrees from 1999. Prior years are/were degrees & minutes (e.g. 45.30 is 45 deg 30 min). It was intended to continue the format conversion back to 1992, but I donât know if this was accomplished. None of the log/slip data prior to 1995 (1986-1994) is loaded. The ZIF database uses decimal degrees with longitude negation.

The year format was originally YY through 1998, thence YYYY. This was thought to have been revised to the YYYY format. However 1996 and 1997 fell through the cracks, retaining the YY format, while 1995 and 1998 were converted to the YYYY format. The format of previous years is presently unknown, as all log/slip tables prior to 1995 are archived.

Original set series in ASCII log data files are maintained. These reflect the keypunch data, which was entered sequentially from the hardcopy logs. So even when otherwise crucial information to discern relative timing of sets is absent, the trail is often easily determined from the sequencing of records in the log data.

The following SQL scripts illustrate set rollup processing, . The examples focus on 2002 because it is the one year we have all 3 main sources of commercial fisheries data - ZIF, log & slip tables, and MARFIS.

#Illustrates set rollup processing from log to ZIF tables.  
#log data  
set pagesize 250;  
set linesize 120;  
select h.log\_code logcode,unit\_area area,  
t\_nbr\_hours hrs,t\_nbr\_sets sets,effort\_measure eff, t\_latitude lat, t\_longitude lon,  
 date\_fished,t\_nbr\_tows tows,  
substr(t\_species\_id\_est,1,3) spec,  
unit\_of\_measure units,t\_live\_weight livewt  
from cl.log\_header\_2002 h,cl.log\_detail\_2002 d,cl.log\_estimate\_2002 e  
where substr(h.log\_code,1,6) = '005645' and substr(h.log\_code,7,6) = '200209'  
and  substr(t\_species\_id\_est,1,3)='110'                                           
and h.log\_code=d.log\_code  
and h.log\_code=e.log\_code  
and d.estimate\_key=e.estimate\_key;

#ZIF  
select a.cfv\_number cfv,a.trip\_num trip,a.sub\_trip\_num subtrip,a.latitude lat,a.longitude lon,  
a.fish\_days,a.effort\_hrs hrs,a.effort\_count cnt,a.caught\_date,  
a.nafo\_unit\_area area, b.live\_wt from cl.sub\_trips\_2002 a,  
cl.identified\_catches\_2002 b where a.cfv\_number=005645 and to\_char(caught\_date,'mm')=09  
and b.species\_code=110 and a.cfv\_number=b.cfv\_number and a.trip\_num=b.trip\_num  
and a.sub\_trip\_num=b.sub\_trip\_num;

#MARFIS  
create table xxx as  
select \*  
 from MFD\_OBFMI.MARFIS\_PRO\_SPC\_INFO  
 where to\_number(substr(to\_char(date\_fished,'DD'),1,2))=3  
 and to\_number(substr(to\_char(date\_fished,'YYYY'),1,4))=2002  
 and to\_number(substr(to\_char(date\_fished,'MM'),1,2))=9  
 and vr\_number\_fishing=5645  
 and species\_code=110  
order by trip\_id,LOG\_EFRT\_STD\_INFO\_ID,date\_fished;

The scripting in [ZIFLogMARFIS.txt](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) (**Log Data Recovery**) demonstrates linking of ZIF and log/slip tables to recover lost information. The objective is to create one year of catch/effort data to conduct CPUE anallyses for the particular gears and species addressed. It does not provide a systematic database-wide matchup, and deliberately tosses records with missing or dubious effort-related data. But it gives the basic idea.

**processing**

Validation and processing steps on regional datasets by HQ:

1. Remove all freshwater landings from the Maritimes files and farmed data from Gulf.

2. Ensuring that all species, gear and port codes have valid labels.

3. Ensure that all the species codes are in the correct species group;  same for vessel length (vessel length group); vessel tonnage; etc

4. Search for duplicate landings between regions: same vessel, species landed, port landed, date landed and weight.

5. Other irregularities often appear when performing adhoc analysis.

**Recent Documentation Review Notes**

**COMPANY** (company) The company which operates the vessel.  ZIFF only maintains the codes, not the names.  The intention is to include all vessels of comÂ­panies which have vessels under individual quota management. Vessel and Buyer Company Codes are assigned in blocks as follows:

                        1          â           1999:   Assigned to Newfoundland

                        2000    â           3999:   Assigned to Gulf

                        4000    â           5999;   Assigned to Maritimes

                        6000    â           6999:   Assigned to Quebec

                        7000    â           9999:   Unassigned

This code will be the key element for referencing a âCompany Tableâ (see âadditional filesâ) that will contain the âCompany Codeâ and âCompany Nameâ, amongst others.

This field is only used by Newfoundland and Maritimes regions.  Other regions should update this field.  Quebec includes a company table with each ZIFF, but places its company code in the last four digits of the FIN (beginning with â40000â).  HQ will ensure that only valid company codes that match those in the Company table are included.

**Review Notes:** The company codes were provided by every region exception Newfoundland from 1986 to 1997, and then by every region (including Newfoundland) from 1998 to 2002. during 2003-2004 only Newfoundland and Maritimes Region provided this data, and during 2005-2006 on ly Newfoundland provided company codes. In 2007 Newfoundland and Maritimes Region provided the codes, the other regions did not.

**QUOTA** (quota): Newfoundland supplies the actual quota ID number whereas Quebec supplies the information as follows:

(position; use)

1â3    species code

4â5    stock area code

6â7    a qualifier to designate particular categories for a given         stock and area.

Whatever the structure (definition) of this code, it must link to the regional âQuota tableâ and provide us with other information such as species, gear and area.

**Review Notes:** This data was provided by every region except the Gulf from 1986 to 1990. Between 1991 and 2002 only Newfoundland and Maritimes Region provided quota data, with the singular exception of year 2000, when the Gulf Region provided it. From 2003 to 2006 only Newfoundland provided this information. In 2007 only Newfoundland and Quebec provided it.

**SARA** (f2.0) (sara\_flag):  This flag will identify species at risk from the species caught or a combination of species and NAFO area.  Given that these species are supposed to be returned to the water, they will not appear on many landing records.  However, some regions are keeping a separate log of the by-catch of these species and we can discuss possible uses for this log at a later date. It is not considered mandatory.  Inform HQ if not available.

**Review Notes:** The SARA variable does not exist in 2007, and may have been dropped. It was originally conceived as a binary variable, but the SARA listing is a multi-level categorization, and does include species that can be landed (special concern). And perhaps some implicit categories would need to be included (e.g. not reviewed, reviewed and okay).

**Creation of a Continuous ZIF Database**

Steps and Objectives

 1. Use the 2007 definition as a template.

 The 2007 definition subsumes the contents of previous ZIF data. No information is to be lost from previous years.

 2. Make the 2003-2006 data correspond to the 2007 definition.

This is just minor formatting, column location adjustments, and removal of some variables left in from processing at NHQ. The processing of the  2003-2006 data was ad hoc as the new ZIF was being designed, such that virtually every year differs in content and format.

 3. Reconcile 1986-2002 data to the 2007 definition.

 The true ZIF was a single ASCII file for each year and region. The database Population Ecology Division (PED) staff typically refer to as ZIF is actually a form of commercial landings database, conceptually approximate to MARFIS in intent, designed by Marine Fish Division (now part of PED). It comprises four types of tables, three of which represent the data that would be included in an actual ZIF file. This will be referred to as the MFD ZIF in this documentation. The fourth separate table is the Catchers table, and this still exists with the new design, although its format and contents were not part of the ZIF redesign process, and currently exists as independent tables for each region.

It had originally been intended to migrate the MFD ZIF to the ZIF2 database, largely an exercise in standardizing codes. The section below titled MFD Historical ZIF (1986-2002) documents this process.

In 2009 it was discovered that the landings represented in the MFD ZIF could be appreciably less than reality. This reflected updates to the ZIF conducted in Ottawa that were not repeated in Maritimes Region. Part of this could be attributed to a belief that any major revisions would be affected within a year or two of the initial creation of a ZIF file for a given year. Turns out that was a misconception. This could have implications for past assessments and investigations that relied on ZIF data. Thus it was decided to use the NHQ ZIF to populate the database for 1986-2002 instead of the MFD ZIF. During preliminary processing of the NHQ ZIF it was discovered that the NHQ ZIF had undergone some secondary processing sometime after 1999, some of which had negative implications for data quality from a Science perspective (NHQ was primarily concerned with economic data). As well, the standardization rationale applied to the NHQ ZIF only applied to species, gears, and length class codes, not several other variables that were resolved with respect to migrating the MFD ZIF to the ZIF2 database. So the documentation for the MFD ZIF migration remains valid in the context of the NHQ ZIF migration. The section below titled NHQ Historical ZIF (1986-2002) documents the processing of this data, with references to the documentation section for the MFD Historical ZIF (1986-2002) where it holds true across sources.

 Resolving the 1986-2002 MFD ZIF data required processing each yearly set of 3 tables into the structure and format of the 2007 definition. Originally no information was to be lost, but we encountered problems with this objective when we switched to the NHQ ZIF. We lost LAND\_STAT\_DIST\_CODE, LAND\_COMMUNITY\_CODE, PLANT\_PROV\_CODE, and PLANT\_COMMUNITY\_CODE. These were based on lookup tables to match against codes of the appropriate NHQ ZIF variables. We didn't lose the information per se, but the NHQ variables might be unfamiliar. For details see MFD Historical ZIF (1986-2002). We lost SUB\_TRIP\_NUM, CATCHERS\_TYPE and TRIP\_SUB\_TRIP\_FLAG. NHQ had no information about these variables. Possibly they were derived locally by processing the ZIF data (???), in which case the source information would not have been lost. The GULFFLAG variable was processed into a binary code variable, where previously it comprised a number of codes. As discussed in an earlier section, nobody seemed able to interpret the old codes, and the new binary coding was deliberated to reflect what we know today (maybe not the best rationale for reducing information content). The SPQUAL (species quality) code variable was recently discarded by NHQ as it was considered dubious (unfortunate, it may have been more credible in earlier years for some regions). Thus we lose the variable with respect to 1986-2002, have it for 2003-2008, and maybe will not see it post-2008. And we lost original longline hook count data for Maritimes region in 1986-1987 and Newfoundland region in 1995-1996. Note that lost information can still be obtained from the MFD ZIF, perhaps with some linkage hurdles due to updates and edits applicable only to the NHQ ZIF.

The 2007 Template

The 2007 data provided by NHQ was truthed against the format definition. Some divergences, mostly resulting from format coercion during data input using SPSS, were amended to match the defined format. Other decisions were made regarding inconsistencies between the text (ZIF 2007.doc) and table (ZIF 2007 General Record Layout.xls) documenting the format definition. Most of the results of this exercise were applied to the 2009 renditions of 2006-2008 data. Only those observations remaining potentially relevant in 2009 are included here. Virtually all of these only concern discrepancies between the data and the documentation, and in all cases the data is fine, the documentation just wants for some updating.

spqual â A categorical A1 code variable. The ZIFF 2007 General Record Layout. xls format/definition table declares an F1 format. The F1 format contradicts both the text definition and the data file.

latitude/longitude â A continuous F12.8 variable. The textual documentation calls for a continuous Fn.8 variable, the number of whole number digits not explicitly defined, and the 2003-2008 data files use the F12.8 format. The ZIFF 2007 General Record Layout. xls table format is F6.4. Probably the sign of the longitude was overlooked, given we can usually represent maps in Northwest Atlantic waters without considering it. Similarly the number of whole number digits, appropriate for local waters, will not work for other regions to be included in the ZIF (Pacific, Central & Arctic). The F12.8 format covers all the possibilities using the defined precision.

cfv\_flag, fin\_flag, fin\_type, cfv\_type â All F2 code variables. The text documentation does not stipulate the format, but the codes themselves are defined, none exceed 2 digits, and F2 was used for 2006-2008 data. The ZIFF 2007 General Record Layout. xls table indicates they are all F8, likely the result of copying a default numeric format for input in SPSS.

homeprov â Included in the 2003-2008 data, but not mentioned in the documentation. This will be retained, using the 2007 format of F2 applied to other province code variables.

Interim Years (2003-2006)

2006

 The 2006 data was regenerated after reconciling the 2007 template data, and conforms to the current (2009) ZIF format

2005

gulfflag F2, change to F1

gearcode F2, change to F3

amtgear F4, change to F7

licence F8, change to A12

create null fisharea A8 variable

rename/change setnumbe F4 to seqnum F8

depth F5, change to F8

homeprov F1, change to F2, and relocate after prov

filter\_f delete records = 1 then delete variable

v58\_a delete records = 1 then delete variable

deprange delete

create null F2 variables for fin\_type, vessel\_type

recode fin\_flag A1 C/I/O/Y to F2 2/1/blank/1

recode cfv\_flag A1 Y/N to F2 1/2

2004

gulfflag F2, change to F1

gearcode F2, change to F3

amtgear F4, change to F7

licence F8, change to A12

create null fisharea A8 variable

rename/change setnumbe F4 to seqnum F8

depth F5, change to F8

homeprov F1, change to F2, and relocate after prov

filter\_f delete records = 1 then delete variable

v59\_a delete records = 1 then delete variable

deprange delete

create null F2 variables for fin\_type, vessel\_type

recode fin\_flag A1 C/I/O/Y to F2 2/1/blank/1

recode cfv\_flag A1 Y/N to F2 1/2

2003

gearcode F2, change to F3

amtgear F4, change to F7

create null fisharea A8 variable

rename/change setnum F4 to seqnum F8

depth F5, change to F8

homeprov F8, change to F2, and relocate after prov

filter\_f delete

v49\_a delete

deprange delete

create null F2 variables for vessel\_flag, vessel\_type

fin\_flag F8, change to F2 (all nulls, no coding)

cfv\_flag F8, change to F2 (all nulls, no coding)

MFD Historical ZIF (1986-2002)

The MFD database uses character formats on almost all variables unless they hold dates or numbers with decimals. Where an old variable is retained (does not exist in the revised definition), it is formatted according to its content (numeric or character data).

In the documentation below only those variables requiring attention beyond simple formatting and naming are addressed. Fuller details can be discerned from the [SQL scripting](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) to conduct the merger of historical and recent ZIF data. The bottom of the SQL scripting is the formatting control to use if spooling the data to an ASCII file according to the format definition.

During the process of conducting the merger any encounters with non-systematic errors of a typographical nature were ignored, not fixed or documented.

Where a variable name differs between the Excel table definition and the SPSS datafile, the definition is used. The differences result from truncation of variable names to 8 characters by SPSS.

The names of variables lat and long have been expanded to latitude and longitude. The term âlongâ is a reserved word in SQL.

SQL doesnât like entirely nulled columns (empty variables) when using select clauses to create tables, and 1986-2002 are years of data with no equivalents to many of the variables defined for the new ZIF. Zero-filling is used for homeport, hpfed, gclass (discussed below), portland, plfed, buyerid, buyerport, spgrp (discussed below), amount, units, seqnum, depth, homeprov, fin\_flag, and fin\_type. Blank-filling is used for efflag, fin, licence, and fisharea. These can be nulled post-merger (once the other years provide some actual values for these variables).

The gclass and spgrp variables in post-2002 ZIF data do not exist in the MFD/PED ZIFish database. Given the issues with gear codes in historical ZIF data between years and regions, these variables are being zero-filled. Both are derived by recoding variables already included in the data. But in both cases historical coding issues would need to be investigated, and may not warrant the effort. An approximate coding for gclass, derived from the ZIF 2007.doc, is illustrated here for future reference but not used in the SQL script to migrate the historical data to the current format. This would only address the mixes of zeroes and nulls for gear codes that the ZIF 2007.doc coding did not address, not differences in gear codes between regions or valid codes that no longer occur.

decode(gearcode,1,2,9,2,10,2,11,2,12,2,13,2,14,2,15,2,16,2,17,2,18,2,19,2,20,2,21,2,22,2,31,2,32,2,33,2,54,2,71,2,74,2,93,2,90,0,99,0,NULL,0,0,0,1) gclass,

**Bold font** flags assumptions, intended actions, or issues in need of resolution. These are items that either warrant review or require feedback to proceed.

catcher\_type (vessel\_type)

recode ? to blank

recode A to 60

recode B to 61

recode F to 93

recode V to blank, make cfv\_flag = 1

recode W to 62

recode M to [mackerel trap; exists from 1997] 10

recode H to [herring weir; exists from 1998, W disappears] 62

**recode P to [pelagic bait; exists from 1998] \_\_[-9 for now]**

sub\_trip\_num [retaining old variable]

**Some similarities to seqnum but differs in some critical respects, and applies to all regions. Concerned that putting its contents into seqnum might lead to confusion, so keeping it separate.**

land\_prov\_code (prov)

**recode 0 to 10 [exists 1998-1999 & 2002, Nfld only]**

recode 1 to 12

recode 2 to 13

recode 3 to 11

recode 4 to 24

recode 5 to 10

**recode 7 to 10 [exists from 1993, Nfld only, guessing** **Labrador]**

land\_stat\_dist\_code [retaining old variable]

land\_community\_code [retaining old variable]

**[Potential Project] These variables are not part of the revised ZIF, which focuses on electoral district and port codes. Electoral districts are not in the MFD representation of ZIF, which focuses on statistical district and community codes. The electoral district exists in a COMMUNITY lookup table, which holds all three codes in parallel. However this table is characterized by multiple possibilities for electoral district. I suspect much or all of this is an artifact of using the ZIF data itself to populate the lookup table. The bulk of the redundancies are likely null codes for electoral district, resulting in 0 or 9 along with a valid code option (and variation in null coding with region, year). This may be complicated by data entry errors producing multiple (usually just 2) valid codes for the odd statistical-community pair. Further complicating matters, the coding systems in use for a given region can change over time, suggesting that the electoral district codes may not be compatible with those in use now (in the systematic sense of the overall coding scheme, recognizing that occasional community assignment to electoral districts can change). This is accounted for in the lookup table by maintaining year-specific codes. Noting discontinuities in the years covered, it may be that the table was only updated when it was known that a coding system changed. If so any missing years would be coded according to the last existing year, although the initial gap for 1986-1989 seems suspicious. So conceivably the lookup table would contain only unique possibilities if based on the underlying coding systems rather than the data itself. But the actual code systems might not be equivalent over time (beyond changes in the electoral districts themselves). If we want electoral districts for historical data, we will need to truth out all the above assumptions, construct an appropriate replacement for the lookup table, and presumably standardize the codes to whatever system or systems (regional variation) are currently in use.**

**estimated\_weight [old variable, not retained; log estimate of weight 1991-1998,** **Quebec** **only]**

**disc\_weight [old variable, not retained; log estimate of discard weight, no data at all]**

size\_code (spsize)

**recode 7 to \_\_ [left as is, no overlap][Nfld/1998 only; then no size codes from Nfld 1999-2002]**

quality\_code (spqual)

**No recoding, left as is.**

**[Potential Project] The ZIF 2007.doc notes extra codes for Gulf data, but they may originate with Maritimes data. The code systems evolved over time, with Maritimes the first to âexpandâ them. Later Nfld adopted a numeric code system. Etc. Below I go through some motions of recoding, but do not actually intend to do this. The codes will be left as is, and defined as region specific and changing over time. Possibly these should be rationalized.**

**add code J = Junk [from 1988, Maritimes only]**

**add code O = Offal [from 1988, Maritimes only]**

**add code M = \_\_\_\_ [from 1988, Maritimes only]**

**add code P = Premium [from 1989, Maritimes only, from 1999** **Quebec]**

**add code Z = \_\_\_\_ [from 1989, Maritimes only, from 1991** **Quebec]**

**recode 0-9 = \_\_\_\_\_ [from 1999, Nfld only]**

**recode \_\_\_\_ = \_\_\_\_\_[just 2000, Gulf only, custom codes EFHNSTUWZ**

form\_code (spform)

**No recoding, left as is.**

**[Potential Project]. Same issue as dealt with in achieving the 2007 ZIF. Exactly what do the codes mean per region and year, and what were the conversion factors?**

unit\_code (spmeas)

**recode N to \_\_ [left as is, nooverlap][from 1998, Nfld only]**

stock\_area\_code, quota\_qualifier

**Incorporated into âquotaâ variable (species code + stock\_area\_code + quota\_qualifier), taken as is.**

**[Potential Project] We donât have lookup tables to make sense of these historically. Appropriate lookup table definitions exist, but the tables are empty. As these codes are specific by year and region, they are simply processed as is. Gross differences in coding schemes at both multiregional and regional levels can be discerned for different periods in time.**

plant\_prov\_code [retaining old variable]

**recode 0 to blank**

recode 1 to 12

recode 2 to 13

recode 3 to 11

recode 4 to 24

recode 5 to 10

**recode 6 to \_\_ [left as is, no overlap][from 1999, Nfld only]**

**recode 9 to \_\_ [left as is, no overlap ][from 1994,** **Quebec** **only]**

plant\_stat\_dist\_code [retaining old variable], plant\_community\_code [retaining old variable]

**See earlier comments related to land\_stat\_dist and land\_community\_code.**

**level\_of\_log\_info [old variable, not retained][Quebec-only 1991-1998]**

**gear\_code [old variable, not retained]**

**Not to be confused with the revised ZIF gearcode varaible, which equals the 1986-2002 ZIF gear\_type variable, gear\_code was a Quebec-only variable used from 1991 to 1998 to provide details associated with dredge gears.**

position\_qualifier [retaining old variable]

**No recoding, taken as is.**

**Codes 0-3, 0 likely null. No definition for these codes, donât know if all regions using compatible codes or any changes over time.**

**map\_no [old variable, not retained][ Quebec-only 1991-1998]**

**grid\_lat, grid\_long [old variables, not retained]**

**Gulf (1991-2002) and** **Quebec** **(1991-1998) used a grid square map to assign lats/longs. Assuming the position\_qualifier will suffice if we need background on lat/long derivation.**

trip\_sub\_trip\_flag [retaining old variable]

**This variable documents how the subtrip records were identified. Code 00 does not exist due to the subtrip processing problem discussed in Section II. The intended 00 records became 02 records.**

**00 Just another row of a subtrip**

**11 The first record of the sorted input file**

**21 The first row of a new trip based on a change in CFV**

**31 The first row of a new trip based on a change in trip number**

**41 The first row of a new trip based on a change in landing date**

**02 The first row of a subtrip based on the effort flag**

**03 The first row of a subtrip based on a change in an effort field**

**hull\_type\_code, hold\_capacity, year\_built, name, location\_description [old variables, not retained]**

**From the catchers tables we only extract the tonnage, length and horsepower code/value paired variables expected to be included in ZIF files. All variables listed above except name are null throughout in any case. The name can be retrieved from the catcher tables as required, and as has been the practice to date.**

SQL Scripting

The [scripting](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation) to create a multiregional ZIF file for a given year in 1986-2002, compatible with the ZIF format as revised in 2007, is primarily concerned with variable contents, names and formats. Noting that problem codes for variables intended to be recoded do conflict with new codes, the beta version of the migrated 2002 ZIF data just recodes where known, leaves the old code where not known.

Following the initial creation of the 1986-2008 ZIF database accroding to the 2007 format, a number of standardizations were conducted to ensure that the database used consistent coding and content across all regions and years. Especially relevant are inclusion the results of a major inter-regional reconciliation exercise concerned with species codes, which had 'evolved' separately in each region to a greater extent than some imagined. Depth codes had also diverged in some very unfortunate ways, such that much depth data may only be recoverable through the source region, but the codes have been standardized to properly reflect the situation. Full details and SQL scripting of these and other standardizations can be found in [post-load processing](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation).

NHQ Historical ZIF (1986-2002)

Most of the documentation pertaining to the MFD ZIF is equally applicable to the NHQ ZIF. The exercise of migrating the NHQ ZIF to the ZIF2 database was largely just applying much of the same processing. This section focuses on issues unique to the NHQ ZIF.

The NHQ ZIF data was loaded to Oracle using LoadNHQZIF.sql, editing the names of the input ASCII data files and their target Oracle tables for each titular year.

Date fields DATELAND and CTCHDATE were changed from formally date-formatted fields to numeric fields in the database to preserve year/month information in cases where the day was 00. SQL\*Loader regarded such dates as nulls. This might explain why these dates in the MFD ZIF were missing (null). This will also result in some cases of impossible dates (e.g. Feb 29) that were also represented by null dates in the MFD ZIF.

The YEAR variable holds the titular year of the source ZIF datafile. Previously (MFD ZIF) it contained the actual year of landing, such that a small number of records might be the year following the titular year. These are records associated with trips that started in one year but ended the next year (e.g. Nov 1989 to Dec 1990). Thus the YEAR variable identifies the source datafile.

The Maritimes Region component of the 1987 ZIF had corrupted values for calculated effort variables (TRFRAC, DAYSEA, DAYSGR, DAYSFISH, HOURFISH). These were fixed by splitting the load for 1987 into two jobs, Load1987A.sql for non-Maritimes data and Load1987B.sql for Maritmes data.

The initial loads of NHQ data included a filter variable used by NHQ to take out aquaculture and freshwater records from the ZIF. It is usually applied and removed at NHQ so we don't see it. But we were not sure if it had been applied to all the datasets provided, so left it in as a precaution. Turns out it was not applied to the 2002 data,  thus we applied it post-load as an update. The NHQFILTER variable was subsequently dropped from all tables.

The bogus CFV 'B00000', which appears to have been recoded to '000000' in the MFD ZIF, is left alone. The NHQ ZIF has both 'B00000' and '000000', so maybe the 'B' means something.

Recoded the effort flag from -/blank to Y/N as necessary (see MFD ZIF documentation).

The species code standardization by NHQ missed wolffish (didn't realize catfish=wolffish, plus Newfoundland misunderstood some of the coding protocol). Thus needed to recode species code 174 to 194 (old and new code for unspecified wolffish) and 187 to 191 (Nfld and new code for Northern Wolffish).

Standardized the depth codes (see MFD ZIF documentation).

Recoded province to meet the new definition (see MFD ZIF documentation).

Final Steps

Each of the 1986-2008 tables was processed into the final format and appended together as a single multi-year table. It is hoped this will prove more convenient than the three-tables-per-year layout of the MFD ZIF. The SQL scripting for creating the new database resides in [CreateZIF1986\_2008.sql](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation). It is not a batch job, but requires a separate run for every year, usually with editing related to which year is being processed. It can be revised to function as a batch job if/when we feel current ZIF processing has stabilized.

A conceptually simpler table than the current ZIF could be derived by treating the species, currently resolved as individual records, as columns. This would reduce requirements of users of the data to understand how to distinguish record membership at the set/subtrip/trip/etc level. It is also more efficient to create this form of table in terms of computer resources. However it is presently unknown how the species by record versus species by column table structures would compare with respect to access and storage efficiencies. The scripting to produce the CPUE database [Log Data Recovery] uses the species column approach, and was more efficient than the species record approach for this purpose. But it was applied to subsets of fisheries data, focusing on species and gears associated with analyses of commercial catch rates, not every species and gear in the database.  Ramifications of the species column approach would need to be investigated before proceeding.

Processing of ZIF data since 2002 underwent some unexpected changes, variously deliberate and inadvertant, that are described here. These changes can affect earlier years than 2002, as some systematic revisions across all years were conducted. Some of these revisions likely involved re-acquiring ZIF data from the regions for previous years, and aspects of data processing in later years varied from historical periods. Trip numbers disappeared for 2002-2005 Maritimes Region data, Newfoundland Region since 1998, and the Gulf Region in 2006. These are blanket omissions for entire years, not to be confused with valid missing trips that occur when insufficient information is associated with landings to acertain trips.

Anecdotal information in past years suggested that the EFFLAG variable, intended to flag one record of a related group of records as the umbrella record for effort data (e.g. the individual records for each species in a set), was generally corrupted. However no actual documentation of this situation existed, such that we still had the intended expectation from its definition. It was straightforward to truth the EFFLAG variable against identified set records for regions with unaggregated data processing. It is indeed incorrectly determined in Maritimes and Newfoundland Region data, and perhaps should be assumed to be unreliable for Gulf and Quebec regions (aggregated data, EFFLAG cannot be truthed). It may have been created using incomplete subtrip definitions. Subtrip, though defined in the context of processing ZIF, was not actually a variable in ZIF data. The definition itself underwent some evolution over the years, and its applicability to the processing to derive EFFLAG may have been overlooked.

NAFO Division/Subarea/Statistical Unit Area coding is not standardized. We see various protocols among regions, and these may differ within regions between time periods. Examples are upper versus lower-case unit areas, and right versus left justification of unit areas (occurs in 3rd or 4th position of format).

**Project List (noted if done)**

**other regions**

It would be worthwhile to obtain historical documentation from the other regions, along the same lines as we have for Maritimes Region. And possibly Jim McMillan has more documentation for Maritimes Region than is provided here. This would be especially helpful for the 1986-1994 period, poorly addressed by this document. Pre-1995 subtrip definitions are a crucial gap. It might also make sense of the depth coding used by Newfoundland since 1995.

**historical ZIF effort variables**

Might be good to provide the historical series of definitions for the 3 effort variables days fished/at sea/on ground. This might only be worthwhile if we can get definitions from all regions. The final stages of calculation for Maritimes is provided by **C0130CE.doc (**[**historical ZIF catch-effort variables**](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation)**)** , but this does not include initial determinations of the variables prior to standardization on trip\_fraction. We would also want to know when the software being documented was implemented, and what was done for previous years.

**species codes (done)**

Might want to reconcile the species codes with current rules. Differences between regions and over time within region could cause confusion. Clouding the issue, some species may not have been covered by DMP all the way back to 1986, so have to distinguish coding protocols from starting years of monitoring.

**unspecified**

The same type of proration by DMP contractors may have been applied to other species than flatfish. This should be investigated, at least with respect to SARA candidates (e.g. wolffish, skates).

**unknown depths**

Might want to simply add an estimated depth variable to the database. Would need to obtain and include geographic matrices of depths for Quebec and Newfoundland Regions. Maybe Jerry already has this? Some background:

ZIFF records without depth information but with latitude/longitude coordinates can be provided an estimated depth based on various sources of depth contouring data. Geographic matrices of depths were originally computed to detail the depth compositions of Research Vessel Survey Strata. This exercise was conducted for all Maritimes Region surveys, as well as that of the Southern Gulf of St Lawrence. A derivation of the resulting **[DepthGrid](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/supporting-zif-documentation" \o "Supporting ZIF documentation)**table for use with commercial fisheries data was created as table DGZ. This table can be linked to any ZIFF data with geographic coordinates that comes from Canadian waters west of Newfoundland and south of the Northern Gulf of St Lawrence. Thus depths can be obtained for much of the data if missing.

#Example of using a depth grid to assign depths to records with unknown depth zone (zifmar is a completed multiyear database in which annual

#files, such as that created by the example [log and slip tables](http://vsnsbiovdc.ent.dfo-mpo.ca/maritimes/commercial-landings/zif-1/zif-update-2009/Plone/data/div-162/zif-1/zif-update-2009/status-1986-2002-zif-procedures/#log-and-slip-tables), are merged).

drop table zzz;  
create table zzz as  
select  
yland,nafo\_division\_code,nafo\_unit\_area,  
gear\_type,main\_species\_caught,  
decode(z.depthzone,'10',d.depthzone,z.depthzone) depthzone,  
z.latitude,z.longitude,  
msc\_tons  
from zifmar z, dgz d  
where  
(trunc(z.latitude,1)+decode((round(z.latitude,2)-trunc(z.latitude,1)),.00,.00,.01,.00,.02,.02,.03,.03,.04,.03,.05,.05,.06,.05,.07,.07,.08,.08,.09,.08,null)) = d.latitude (+) and  
-(trunc(z.longitude,1)+decode((round(z.longitude,2)-trunc(z.longitude,1)),.00,.00,.01,.00,.02,.02,.03,.03,.04,.03,.05,.05,.06,.05,.07,.07,.08,.08,.09,.08,null)) = d.longitude (+);

#Haddock example for depth reconciliation  
#Before substituting depths  
select decode(latitude,0,0,null,0,1) location, yland, depthzone, sum(msc\_tons) tons from zzz  
where  
main\_species\_caught='110'  
group by decode(latitude,0,0,null,0,1),yland,depthzone  
order by decode(latitude,0,0,null,0,1),yland,depthzone;

#Depth zones (DZ) according to MARFIS protocol (0=very deep, 10=unknown), including for ZIF years.  
NOLOCATION      YLAND DZ       TONS  
---------- ---------- -- ----------  
         0       1995 10 43.1846511  
         0       1996 10 173.979464  
         0       1997 1  .642334587  
         0       1997 10 236.038233  
         0       1998 10 308.381476  
         0       1999 10 196.120667  
         0       2000 10 336.584519  
         0       2000 3    1.001541  
         0       2001 10  322.00945  
         0       2002 10 316.725112  
         0       2003 10 230.008692  
         0       2004 10 102.737619  
         0       2005 10  80.238604

  LOCATION      YLAND DZ       TONS  
---------- ---------- -- ----------  
         1       1995 10 4431.30364  
         1       1995 1  47.1037891  
         1       1995 2   120.79208  
         1       1995 3   82.681797  
         1       1995 4  .742200985  
         1       1995 5  .172001857  
         1       1996 10 6295.62217  
         1       1996 1  29.1122306  
         1       1996 2  266.643162  
         1       1996 3  81.2379416  
         1       1996 4  6.09113099  
         1       1996 5  .722868379  
         1       1997 10 5296.56161  
         1       1997 1  15.3441216  
         1       1997 2  63.8689512  
         1       1997 3  86.3304313  
         1       1997 4  3.39883877  
         1       1997 5  2.12853172  
         1       1997 7     .674045  
         1       1997 0  .312371901  
         1       1998 10 7751.65495  
         1       1998 1  18.1223534  
         1       1998 2  83.5359493  
         1       1998 3   75.107505  
         1       1998 4  5.79450208  
         1       1999 10 7475.05408  
         1       1999 1  84.2865032  
         1       1999 2  102.641954  
         1       1999 3  146.668693  
         1       1999 4  22.5255843  
         1       1999 5  1.02122872  
         1       1999 6  1.69375406  
         1       2000 10 9056.58437  
         1       2000 1  309.467427  
         1       2000 2  153.534986  
         1       2000 3  217.042026  
         1       2000 4  34.9811547  
         1       2000 5  .383933463  
         1       2001 10 11498.1503  
         1       2001 1  361.955949  
         1       2001 2   435.75614  
         1       2001 3  111.321795  
         1       2001 4  29.4738052  
         1       2001 5  2.22012824  
         1       2001 6  .580903031  
         1       2001 9    1.732402  
         1       2002 10 11793.9346  
         1       2003 10 12958.9631  
         1       2004 10 14450.3114  
         1       2005 10 17812.8287

#After substituting depths (requires LOCATION)  
select decode(latitude,0,0,null,0,1) location, yland, depthzone, sum(msc\_tons) tons from zifmar  
where  
main\_species\_caught='110'  
group by decode(latitude,0,0,null,0,1),yland,depthzone  
order by decode(latitude,0,0,null,0,1),yland,depthzone;

         1       1995 10 923.481576  
         1       1995 1  59.8932364  
         1       1995 2  1267.16067  
         1       1995 3  1205.87718  
         1       1995 4  724.432204  
         1       1995 5  281.800626  
         1       1995 6  96.6245114  
         1       1995 7  148.805558  
         1       1995 8  3.64495357  
         1       1996 10 2493.62905  
         1       1996 1  59.5205186  
         1       1996 2  1966.75688  
         1       1996 3  1255.47066  
         1       1996 4  326.374584  
         1       1996 5  510.677377  
         1       1996 6  90.1881432  
         1       1996 7  7.25560227  
         1       1996 8  8.18994371  
         1       1996 9  5.51720444  
         1       1997 10 2854.69094  
         1       1997 1  33.8677688  
         1       1997 2  839.544518  
         1       1997 3  789.307268  
         1       1997 4  273.763985  
         1       1997 5  471.223335  
         1       1997 6   209.97399  
         1       1997 7   5.2064131  
         1       1997 8  7.17475569  
         1       1997 0  .312371901  
         1       1998 10  2140.8829  
         1       1998 1  27.8113721  
         1       1998 2  1276.56226  
         1       1998 3  1955.78776  
         1       1998 4  915.691881  
         1       1998 5  723.020703  
         1       1998 6  436.030036  
         1       1998 7  296.209568  
         1       1998 8   148.34239  
         1       1998 9  70.1041854  
         1       1999 10 1852.27993  
         1       1999 1  88.6228612  
         1       1999 2  1418.96247  
         1       1999 3  2202.40883  
         1       1999 4  981.928603  
         1       1999 5  452.999976  
         1       1999 6  452.043062  
         1       1999 7  200.835609  
         1       1999 8  220.925641  
         1       1999 9  .012644939  
         1       2000 10 1704.14461  
         1       2000 1   337.47609  
         1       2000 2  2783.78742  
         1       2000 3  3238.10392  
         1       2000 4  868.566647  
         1       2000 5  456.305464  
         1       2000 6   314.13727  
         1       2000 7  128.546875  
         1       2000 8  57.2758964  
         1       2001 10 2643.58568  
         1       2001 1  380.468127  
         1       2001 2  3198.90997  
         1       2001 3  3447.89877  
         1       2001 4  1265.55108  
         1       2001 5  785.547644  
         1       2001 6   206.33199  
         1       2001 7  361.341487  
         1       2001 8  230.035371  
         1       2001 9    1.732402  
         1       2002 10 3425.16041  
         1       2002 1   13.451182  
         1       2002 2  2889.49103  
         1       2002 3  2861.66753  
         1       2002 4  1140.00271  
         1       2002 5  652.304715  
         1       2002 6  367.692418  
         1       2002 7  303.973022  
         1       2002 8  206.130518  
         1       2002 9     .839944  
         1       2003 10 3504.63232  
         1       2003 1    1.499591  
         1       2003 2   3330.6203  
         1       2003 3  2825.78365  
         1       2003 4  1409.27758  
         1       2003 5   850.18446  
         1       2003 6  393.659937  
         1       2003 7  405.426956  
         1       2003 8   311.54462  
         1       2004 10 1881.36812  
         1       2004 1    1.329981  
         1       2004 2  6936.43868  
         1       2004 3   3217.5079  
         1       2004 4  1219.45776  
         1       2004 5  400.126934  
         1       2004 6  191.341544  
         1       2004 7  427.864498  
         1       2004 8  390.918023  
         1       2004 9    1.014842  
         1       2005 10 4384.80964  
         1       2005 1    1.056296  
         1       2005 2  6545.60047  
         1       2005 3  5181.52545  
         1       2005 4  746.870983  
         1       2005 5  274.804782  
         1       2005 6    216.2457  
         1       2005 7  462.789384  
         1       2005 8  277.157939  
         1       2005 9     2.15434

#How much of the known-position catch gets a depth (just MARFIS years since no prior depth data)  
select decode(latitude,0,0,null,0,1) location, yland, decode(depthzone,'10','10','1') depthdata, sum(msc\_tons) tons from zifmar  
where  
main\_species\_caught='110' and yland>=2002  
group by decode(latitude,0,0,null,0,1),yland,decode(depthzone,'10','10','1')  
order by decode(latitude,0,0,null,0,1),yland,decode(depthzone,'10','10','1');

         1       2002 1  8435.55308  
         1       2002 10 3425.16041  
         1       2003 1  9527.99709  
         1       2003 10 3504.63232  
         1       2004 1  12786.0002  
         1       2004 10 1881.36812  
         1       2005 1  13708.2053  
         1       2005 10 4384.80964

**Minor Task List (*noted if done)***

**effort\_flag *(done)***

All codes for effort\_flag are recoded to the Y/N format during post-load processing. Raises the issue of documentation detail. We need to keep track of all differences between years and regions with respect to structure, coding, formats. But much of the coding and format variation can be eliminated to create the Oracle database. So details like original coding schemes in the ZIFF data received from HQ might be too much clutter for a âuserâ version of the documentation. This is the only easily corrected example of documentation left in the above sections, all other similar situations are given in this section.

**confusion (*done*)**

During post-load processing we standardize all species codes across years and regions to be consistent. The current species codes in the ZIF database are the result of a reconciliation exercise conducted by NHQ to produce a coherent multiregional ZIF with common values for code variables.

**main species caught**

Confirm that Main Species Caught is the dominant species by weight in all years and regions (with the known exception of halibut in Maritimes Region through 2001).

**latitude/longitude *(done*)**

Latitude/longitude formats are standardized to decimal degrees during post-load processing. As well the sign (+/-) of the value now reflects the coordinate scheme for the globe as a whole. Thus longitudes for any of the Atlantic regions are negative numbers.

**Remember:** **Need to recover the Catchers tables for 2003 onwards.**